Dong et al. *Journal of Inequalities and Applications* (2023) 2023:97
https://doi.org/10.1186/s13660-023-03011-2

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH

Open Access

On the unconstrained optimization reformulations for a class of stochastic vector variational inequality problems

Dan-dan Dong[1], Guo-ji Tang[2*] and Hui-ming Qiu[1]

*Correspondence:
guojvtang@126.com
[2]School of Mathematics and Physics, Guangxi Key Laboratory of Hybrid Computation and IC Design Analysis, Guangxi Minzu University, Guangxi, 530006, P.R. China
Full list of author information is available at the end of the article

Abstract

In this paper, a class of stochastic vector variational inequality (SVVI) problems are considered. By employing the idea of a D-gap function, the SVVI problem is reformulated as a deterministic model, which is an unconstrained expected residual minimization (UERM) problem, while it is reformulated as a constrained expected residual minimization problem in the work of Zhao et al. Then, the properties of the objective function are investigated and a sample average approximation approach is proposed for solving the UERM problem. Convergence of the proposed approach for global optimal solutions and stationary points is analyzed. Moreover, we consider another deterministic formulation, i.e., the expected value (EV) formulation for an SVVI problem, and the global error bound of a D-gap function based on the EV formulation is given.

Keywords: Stochastic vector variational inequality; D-gap function; Sample average approximation; Convergence; Error bound

1 Introduction

It is well known that the vector variational inequality (VVI) is an effective tool for studying vector optimization problems [16]. The concept of VVI was originally introduced in finite-dimensional Euclidean spaces by Giannessi [9]. Since then, the VVI problem has been extensively considered in studying some related problems, such as vector equilibrium [10], traffic network equilibrium [6, 18], market equilibrium [2], and so on. Since we often encounter uncertain problems in the real world, it is meaningful to consider the stochastic version of a vector variational inequality.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $K \subseteq \mathbb{R}^n$ be a nonempty, closed and convex set, and $\xi(\omega) : \Omega \to \Xi$ a random vector with supported on a closed set $\Xi \subset \mathbb{R}^r$. The stochastic vector variational inequality (SVVI) problem is to find $x^* \in K$ such that

$$\begin{aligned} &\big((y-x^*)^\top F_1(x^*,\xi(\omega)),\dots,(y-x^*)^\top F_m(x^*,\xi(\omega))\big) \\ &\notin \text{-int}\,\mathbb{R}^m_+, \quad \forall y \in K, \text{a.s. } \xi(\omega) \in \Xi, \end{aligned} \tag{1}$$

where the vector-valued functions $F_j : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$ ($j = 1, \dots, m$) contain certain random variables and "a.s." is the abbreviation for "almost surely" under the given probability measure. For convenience, we denote $\xi(\omega)$ by ξ and $F := (F_1, \dots, F_m)$. It is easy to see that problem (1) includes some models as special cases, for example:

(i) If Ω only contains a realization, then SVVI (1) reduces to a VVI problem, that is, to find a vector $x^* \in K$ such that

$$\left(\left(y-x^*\right)^\top F_1\left(x^*\right), \dots, \left(y-x^*\right)^\top F_m\left(x^*\right)\right) \notin \text{-int}\,\mathbb{R}^m_+, \quad \forall y \in K.$$

The interested reader is referred to the monographs [1, 10] and some papers [3, 4, 13, 16, 17, 27] for more details on VVI problems.

(ii) When $m = 1$, the SVVI (1) reduces to a stochastic variational inequality (SVI) problem, that is, to find a vector $x^* \in K$ such that

$$\left(y-x^*\right)^\top F_1\left(x^*, \xi\right) \geq 0, \quad \forall y \in K, \text{ a.s.} \xi \in \Xi.$$

For more details on SVI problems, please refer to some papers, for example [11, 12, 19, 21–25, 28].

(iii) When Ω only contains a realization and $m = 1$, the SVVI (1) reduces to a classical variational inequality (VI) problem, that is, to find a vector $x^* \in K$ such that

$$\left(y-x^*\right)^\top F_1\left(x^*\right) \geq 0, \quad \forall y \in K. \tag{2}$$

For VI problems, please refer to, for example, the monograph [7, 8].

In the past decades, much attention has been paid to the study of some SVI problems. We would like to point out that an important issue in the field of SVI problems is how to solve them. Gürkan–Özge–Robinson [11] proposed a sample-path solution, also called the expected value (EV) approach, for dealing with the SVI problem. Subsequently, the SVI is turned into a certain VI problem. After that, Jiang–Xu [14] reformulated SVI as an optimization problem based on the EV formulation and studied the global convergence results of the presented approach. He et al. [12] utilized a sample average approximation approach for dealing with the SVI based on the EV formulation. In addition, another approach, called the expected residual minimization (ERM), was presented by Chen–Fukushima [5] for solving a stochastic linear complementarity problem. Luo–Lin [23, 24] applied the ERM approach as a natural extension to solve the SVI problems, where the functions $F(x, \xi)$ are affine and nonlinear, respectively. Later, Ma–Wu–Huang [28] applied the ERM approach to solve a stochastic affine variational inequality problem with nonlinear perturbations based on the work of Luo–Lin [23]. Lu–Li [21] and Lu–Li–Yang [22] introduced a new formulation called weighted expected residual minimization (WERM) for solving the SVI problem based on the works of Ma–Wu–Huang [28] and Luo–Lin [24], respectively.

In the recent literature, there are few studies about SVVI problems [26, 27, 33]. Zhao et al. [33] applied the ERM approach for solving the SVVI problem, which generalized some results of Luo–Lin in [23, 24] from an SVI problem to an SVVI problem. There, the SVVI is converted into a constrained optimization problem. We would like to point out that it is interesting to propose an unconstrained optimization formulation to deal with the SVVI problem. By using the idea of the D-gap function, we propose an unconstrained optimization reformulation for a class of SVVI problems. For more details on the D-gap function, please refer to [12, 15, 20, 32], for instance.

The remainder of this paper is structured as follows: in Sect. 2, we recall some fundamental results that will be used in the following sections. In Sect. 3, we investigate the properties of the objective function θ, i.e., continuous differentiability and the boundedness of the level set. In Sect. 4, we use the SAA approach to approximate the expected value of the objective function, and investigate the convergence of the SAA approach for global optimal solutions and stationary points. In Sect. 5, another deterministic formulation, i.e., the EV formulation for the SVVI problem is considered. And we give the global error bound of the D-gap function $g_{\alpha\beta}(\cdot,\cdot)$ based on the EV formulation.

2 Preliminaries

Recently, Zhao et al. [33, Eq. (6)] presented an equivalent scalar variational inequality (with certain constraints) formulation of the SVVI problem, which is to find $(x^*,\lambda^*) \in K \times \Lambda$ such that

$$\big(y - x^*\big)^\top \sum_{j=1}^{m} \lambda_j^* F_j\big(x^*, \xi\big) \geq 0, \quad \forall y \in K, \text{a.s.} \xi \in \Xi, \tag{3}$$

where $\Lambda := \{\lambda \in \mathbb{R}^m : \lambda_j \geq 0, \sum_{j=1}^{m} \lambda_j = 1\}$.

Generally, the presence of a random variable $\xi \in \Xi$ in problem (3) leads to no solution. It is therefore particularly significant to give a reasonable deterministic reformulation for problem (3).

Following the ideas of Yamashita–Taji–Fukushima [32, Eq. (6)] and Zhao et al. [33, Eq. (7)], we introduce the D-gap function $g_{\alpha\beta} : \mathbb{R}^n \times \Lambda \times \Xi \to [0,\infty)$ for SVVI (3) as follows:

$$g_{\alpha\beta}(x,\lambda,\xi) := g_\alpha(x,\lambda,\xi) - g_\beta(x,\lambda,\xi), \quad x \in \mathbb{R}^n, \lambda \in \Lambda, \text{a.s. } \xi \in \Xi, \tag{4}$$

where $0 < \alpha < \beta$ and $g_\gamma(\gamma = \alpha,\beta) : \mathbb{R}^n \times \Lambda \times \Xi \to \mathbb{R}$ is the regularized gap function originated from [33, Eq. (7)] by Zhao et al., which is defined as

$$g_\gamma(x,\lambda,\xi) := \max_{y\in K} \left\{ (x-y)^\top \sum_{j=1}^{m} \lambda_j F_j(x,\xi) - \frac{\gamma}{2}\|x-y\|^2 \right\}. \tag{5}$$

It is easy to see that

$$g_\gamma(x,\lambda,\xi) = \big(x - H_\gamma(x,\lambda,\xi)\big)^\top \sum_{j=1}^{m} \lambda_j F_j(x,\xi) - \frac{\gamma}{2}\big\|x - H_\gamma(x,\lambda,\xi)\big\|^2, \tag{6}$$

where

$$H_\gamma(x,\lambda,\xi) := \mathrm{Proj}_K\left(x - \gamma^{-1} \sum_{j=1}^{m} \lambda_j F_j(x,\xi) \right). \tag{7}$$

In what follows, we assume that D-gap function $g_{\alpha\beta}(x,\lambda,\cdot)$ is integrable on Ξ for each (x,λ).

Remark 1

(i) If $m = 1$ and $K = \mathbb{R}^n_+$, then the D-gap function $g_{\alpha\beta}$ reduces to the corresponding gap function [20, Eq. (3)] for the stochastic complementarity problem of Liu–Li [20];

(ii) If Ω only involves a realization and $m = 1$, then the D-gap function $g_{\alpha\beta}$ is the same as the gap function [32, Eq. (6)] considered by Yamashita–Taji–Fukushima [32] with $\phi(x,y) = (1/2)\|x-y\|^2$.

In what follows, similar to Liu–Li [20, Eq. (6)], we present an unconstrained ERM (UERM) formulation for problem (3):

$$\min_{x\in\mathbb{R}^n,\lambda\in\Lambda} \theta(x,\lambda) := \mathbb{E}\big[g_{\alpha\beta}(x,\lambda,\xi)\big], \tag{8}$$

where $\mathbb{E}$ denotes the mathematical expectation with respect to the law of $\xi \in \Xi$.

Following the work of Zhao et al. [33, p. 551], we adopt the following assumptions, which will be used in the sequel:

(a) For $\xi \in \Xi$ and each $j = 1,\ldots,m$, the function $F_j(\cdot,\xi)$ is a.s. continuously differentiable on $\mathbb{R}^n$.

(b) There exists an integrable function $\kappa(\xi)$ such that

$$\mathbb{E}\big[\kappa^2(\xi)\big] < +\infty \quad \text{and} \quad \sum_{j=1}^m \big\|F_j(x,\xi)\big\| + \sum_{j=1}^m \big\|\nabla_x F_j(x,\xi)\big\|_{\mathcal{F}} \le \kappa(\xi),$$

hold a.s. for any $x \in \mathbb{R}^n$ and $\xi \in \Xi$. Here the Frobenius norm $\|\cdot\|_{\mathcal{F}}$ is defined by $\|A\|_{\mathcal{F}} = (\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2)^{\frac{1}{2}}$ for a given matrix A.

(c) For each $j = 1,\ldots,m$, the function $F_j(\cdot,\xi)$ is Lipschitz continuous on $\mathbb{R}^n$ with Lipschitz constant $L_j(\xi)$ satisfying $\mathbb{E}[L_j^2(\xi)] < +\infty$, i.e.,

$$\big\|F_j(y,\xi) - F_j(x,\xi)\big\| \le L_j(\xi)\|y-x\|, \quad \forall x,y \in \mathbb{R}^n.$$

And meanwhile, $L = \max_{1\le j\le m} \mathbb{E}[L_j(\xi)]$.

Let us recall some well-known concepts and lemmas, which will be frequently used in the sequel.

Definition 1 Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a mapping.

(i) It is said to to be monotone on $\mathbb{R}^n$ if for any $x,y \in \mathbb{R}^n$,

$$\big(F(y) - F(x)\big)^\top (y-x) \ge 0;$$

(ii) It is said to be strongly monotone on $\mathbb{R}^n$ with modulus $\sigma > 0$ if for any $x,y \in \mathbb{R}^n$,

$$\big(F(y) - F(x)\big)^\top (y-x) \ge \sigma\|y-x\|^2.$$

Definition 2 ([14]) Let $F : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$ be a mapping and $K \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^r$. Then F is said to be uniformly strongly monotone on K with modulus $\mu > 0$ over V, if for almost every $\xi \in V$ and any $x,y \in K$,

$$\big(F(y,\xi) - F(x,\xi)\big)^\top (y-x) \ge \mu\|y-x\|^2.$$

Lemma 1 ([7, Theorem 2.3.3]) *Let $K \subseteq \mathbb{R}^n$ be a nonempty, closed and convex set and let $F : K \to \mathbb{R}^n$ be continuous. If F is strongly monotone on K, then $\mathrm{VI}(K, F)$ has a unique solution.*

Lemma 2 ([8, Theorem 10.2.3]) *Let $K \subseteq \mathbb{R}^n$ be a nonempty, closed and convex set and let $F : K \to \mathbb{R}^n$ be continuous. The following statements are valid for VI (2):*

(a) *For every $x \in K$, $y_c(x) = \Pi_K(x - \alpha^{-1}F(x))$.*
(b) *$\theta_\alpha(x)$ is continuous on K and nonnegative on K.*
(c) *$[\theta_\alpha(x) = 0, x \in K]$ if and only if x solves the VI problem,*

where $\theta_\alpha(x)$ is the regularized gap function for problem (2), which is defined as $\theta_\alpha(x) = \max_{y\in K}\{F(x)^\top(x - y) - \frac{\alpha}{2}\|x - y\|^2\}$.

Lemma 3 ([30, Theorem 16.8]) *Suppose that $f(x, \xi)$ is a measurable and integrable function of ξ for each x in (a, b). Let $\phi(x) = \int f(x, \xi)\mu(d\xi)$. Suppose that for $\xi \in \mathcal{A}$, where $\mathcal{A}$ satisfies $\mu(\Omega - \mathcal{A}) = 0$, $f(x, \xi)$ has in (a, b) a derivative $f'(x, \xi)$; suppose further that $|f'(x, \xi)| \le g(\xi)$ for $\xi \in \mathcal{A}$ and $x \in (a, b)$, where g is integrable. Then $\phi(x)$ has derivative $\int f'(x, \xi)\mu(d\xi)$ on (a, b).*

3 Properties of the objective function θ

In this section, we first investigate the continuous differentiability of the objective function θ. To this end, we give some fundamental lemmas.

Lemma 4 *For any $(x, \lambda) \in \mathbb{R}^n \times \Lambda$ and $\xi \in \Xi$, it holds that*

$$\|x - H_\gamma(x, \lambda, \xi)\| \le \frac{2}{\gamma}\sum_{j=1}^m \|F_j(x, \xi)\|. \tag{9}$$

Proof We have

$$\begin{aligned}\frac{\gamma}{2}\|x - H_\gamma(x, \lambda, \xi)\|^2 &\le \big(x - H_\gamma(x, \lambda, \xi)\big)^\top \sum_{j=1}^m \lambda_j F_j(x, \xi) \\ &\le \|x - H_\gamma(x, \lambda, \xi)\| \sum_{j=1}^m \lambda_j \|F_j(x, \xi)\| \\ &\le \|x - H_\gamma(x, \lambda, \xi)\| \sum_{j=1}^m \|F_j(x, \xi)\|,\end{aligned}$$

where the first inequality follows from the nonnegativity of the regularized gap function g_γ, the second inequality follows from the Cauchy–Schwarz inequality, and the last inequality follows from the fact that $\lambda_j \in [0, 1]$. Thus, we get the desired conclusion. This completes the proof. □

Lemma 5 *Let $0 < \alpha < \beta$. For any $(x, \lambda) \in \mathbb{R}^n \times \Lambda$ and $\xi \in \Xi$, one has*

$$\|H_\beta(x, \lambda, \xi) - H_\alpha(x, \lambda, \xi)\| \le \big(\alpha^{-1} - \beta^{-1}\big)\sum_{j=1}^m \|F_j(x, \xi)\|. \tag{10}$$

Proof It holds that

$$\begin{aligned}
&\left\| H_\beta(x,\lambda,\xi) - H_\alpha(x,\lambda,\xi) \right\| \\
&= \left\| \mathrm{Proj}_K\left(x - \beta^{-1}\sum_{j=1}^m \lambda_j F_j(x,\xi)\right) - \mathrm{Proj}_K\left(x - \alpha^{-1}\sum_{j=1}^m \lambda_j F_j(x,\xi)\right) \right\| \\
&\le \left\| \left(x - \beta^{-1}\sum_{j=1}^m \lambda_j F_j(x,\xi)\right) - \left(x - \alpha^{-1}\sum_{j=1}^m \lambda_j F_j(x,\xi)\right) \right\| \\
&\le \left(\alpha^{-1} - \beta^{-1}\right)\sum_{j=1}^m \left\| F_j(x,\xi)\right\|,
\end{aligned}$$

where the first inequality follows from the nonexpansivity property of the projection operator Proj_K, and the second inequality follows from the triangle inequality of the norm and the fact that $\lambda_j \in [0,1]$. This completes the proof. □

Remark 2 If $m = 1$ and $K = \mathbb{R}^n_+$, then Lemma 5 reduces to [20, Lemma 2.1] due to Liu–Li.

Theorem 1 *Suppose that assumptions (a) and (b) hold. Then, one has*
(i) $g_{\alpha\beta}(\cdot,\cdot,\xi)$ *is continuously differentiable with respect to* (x,λ) *for any* $\xi \in \Xi$;
(ii) $\theta(\cdot,\cdot)$ *is continuously differentiable with respect to* (x,λ) *and*

$$\nabla\theta(x,\lambda) = \mathbb{E}\left[\nabla_{(x,\lambda)} g_{\alpha\beta}(x,\lambda,\xi)\right].$$

Proof (i) Since the functions $F_j(\cdot,\xi)$ $(j = 1,\dots,m)$ are continuously differentiable with respect to x on $\mathbb{R}^n$ by assumption (a), it follows from item (c) of Lemma 2 that the D-gap function $g_{\alpha\beta}(\cdot,\cdot,\xi)$ is continuously differentiable with respect to (x,λ) on $\mathbb{R}^n \times \Lambda$ for any $\xi \in \Xi$.

(ii) From item (c) of Lemma 2, we have

$$\begin{aligned}
&\nabla_{(x,\lambda)} g_{\alpha\beta}(x,\lambda,\xi) \\
&= \begin{pmatrix} \sum_{j=1}^m \lambda_j \nabla_x F_j(x,\xi)(H_\beta(x,\lambda,\xi) - H_\alpha(x,\lambda,\xi)) + \alpha(H_\alpha(x,\lambda,\xi) - x) - \beta(H_\beta(x,\lambda,\xi) - x) \\ (H_\beta(x,\lambda,\xi) - H_\alpha(x,\lambda,\xi))^\top F_1(x,\xi) \\ \vdots \\ (H_\beta(x,\lambda,\xi) - H_\alpha(x,\lambda,\xi))^\top F_m(x,\xi) \end{pmatrix}. \qquad (11)
\end{aligned}$$

Hence, we obtain that

$$\begin{aligned}
&\left\| \nabla_{(x,\lambda)} g_{\alpha\beta}(x,\lambda,\xi) \right\| \\
&\le \left\| \sum_{j=1}^m \lambda_j \nabla_x F_j(x,\xi)\left(H_\beta(x,\lambda,\xi) - H_\alpha(x,\lambda,\xi)\right) + \alpha\left(H_\alpha(x,\lambda,\xi) - x\right) \right. \\
&\quad \left. - \beta\left(H_\beta(x,\lambda,\xi) - x\right) \right\| + \sum_{j=1}^m \left\| F_j(x,\xi)\left(H_\beta(x,\lambda,\xi) - H_\alpha(x,\lambda,\xi)\right) \right\| \qquad (12) \\
&\le \sum_{j=1}^m \left\| \nabla_x F_j(x,\xi)\right\| \left\| H_\beta(x,\lambda,\xi) - H_\alpha(x,\lambda,\xi) \right\| + \alpha \left\| H_\alpha(x,\lambda,\xi) - x \right\|
\end{aligned}$$

$$+\beta\left\|H_\beta(x,\lambda,\xi)-x\right\|+\sum_{j=1}^m\left\|F_j(x,\xi)\right\|\left\|H_\beta(x,\lambda,\xi)-H_\alpha(x,\lambda,\xi)\right\|,$$

where the last inequality follows from the triangle and Cauchy–Schwarz inequalities, as well as the fact that $\lambda_j \in [0,1]$. Taking $\gamma = \alpha, \beta$ in Lemma 4, we get

$$\left\|x-H_\alpha(x,\lambda,\xi)\right\| \le \frac{2}{\alpha}\sum_{j=1}^m\left\|F_j(x,\xi)\right\| \tag{13}$$

and

$$\left\|x-H_\beta(x,\lambda,\xi)\right\| \le \frac{2}{\beta}\sum_{j=1}^m\left\|F_j(x,\xi)\right\|. \tag{14}$$

Thus, we obtain

$$\begin{aligned}
&\left\|\nabla_{(x,\lambda)}g_{\alpha\beta}(x,\lambda,\xi)\right\| \\
&\quad\le \left(\sum_{j=1}^m\left\|\nabla_x F_j(x,\xi)\right\|+\sum_{j=1}^m\left\|F_j(x,\xi)\right\|\right)\left(\left\|H_\beta(x,\lambda,\xi)-H_\alpha(x,\lambda,\xi)\right\|\right) \\
&\qquad+\alpha\left\|H_\alpha(x,\lambda,\xi)-x\right\|+\beta\left\|H_\beta(x,\lambda,\xi)-x\right\| \\
&\quad\le \left(\alpha^{-1}-\beta^{-1}\right)\sum_{j=1}^m\left\|F_j(x,\xi)\right\|\left(\sum_{j=1}^m\left\|\nabla_x F_j(x,\xi)\right\|+\sum_{j=1}^m\left\|F_j(x,\xi)\right\|\right)+4\sum_{j=1}^m\left\|F_j(x,\xi)\right\| \\
&\quad= \left(4+\left(\alpha^{-1}-\beta^{-1}\right)\left(\sum_{j=1}^m\left\|\nabla_x F_j(x,\xi)\right\|+\sum_{j=1}^m\left\|F_j(x,\xi)\right\|\right)\right)\sum_{j=1}^m\left\|F_j(x,\xi)\right\|,
\end{aligned}$$

where the first inequality follows from (12) and the second inequality follows from (13), (14), and (10) in Lemma 5. From Lemma 3, we know that the function θ is continuously differentiable, and simultaneously get that $\nabla\theta(x,\lambda) = \mathbb{E}[\nabla_{(x,\lambda)}g_{\alpha\beta}(x,\lambda,\xi)]$. □

Next, we investigate the boundedness of the level set of the objective function θ. The level set of θ is defined by

$$L_\theta(c) := \left\{(x,\lambda)\in\mathbb{R}^n\times\Lambda : \theta(x,\lambda)\le c\right\}.$$

Lemma 6 *Let* $0<\alpha<\beta$. *For any* $(x,\lambda)\in\mathbb{R}^n\times\Lambda$ *and* $\xi\in\Xi$, *one has*

$$\frac{\beta-\alpha}{2}\left\|x-H_\beta(x,\lambda,\xi)\right\|^2 \le g_{\alpha\beta}(x,\lambda,\xi) \le \frac{\beta-\alpha}{2}\left\|x-H_\alpha(x,\lambda,\xi)\right\|^2. \tag{15}$$

Proof We only prove the first inequality of (15). By the definition of D-gap function $g_{\alpha\beta}$ and Lemma 2, we have

$$g_{\alpha\beta}(x,\lambda,\xi) = \left(x-H_\alpha(x,\lambda,\xi)\right)^\top\sum_{j=1}^m\lambda_j F_j(x,\xi)-\frac{\alpha}{2}\left\|x-H_\alpha(x,\lambda,\xi)\right\|^2$$

$$
\begin{aligned}
&- \big(x - H_\beta(x,\lambda,\xi)\big)^\top \sum_{j=1}^{m} \lambda_j F_j(x,\xi) + \frac{\beta}{2}\big\|x - H_\beta(x,\lambda,\xi)\big\|^2 \\
\geq{}& \big(x - H_\beta(x,\lambda,\xi)\big)^\top \sum_{j=1}^{m} \lambda_j F_j(x,\xi) - \frac{\alpha}{2}\big\|x - H_\beta(x,\lambda,\xi)\big\|^2 \\
&- \big(x - H_\beta(x,\lambda,\xi)\big)^\top \sum_{j=1}^{m} \lambda_j F_j(x,\xi) + \frac{\beta}{2}\big\|x - H_\beta(x,\lambda,\xi)\big\|^2 \\
={}& \frac{\beta-\alpha}{2}\big\|x - H_\beta(x,\lambda,\xi)\big\|^2.
\end{aligned}
$$

In a similar way, we obtain the second inequality of (15). This completes the proof. □

Theorem 2 *If K is a compact set, then, for any $c \geq 0$, the level set $L_\theta(c)$ of the objective function θ is bounded.*

Proof Suppose on the contrary that there exists a $\bar{c} \geq 0$ such that $L_\theta(\bar{c})$ is unbounded. This implies that there exists a sequence $\{(x^k,\lambda^k)\} \subset L_\theta(\bar{c})$ such that

$$\lim_{k\to\infty} \big\|\big(x^k,\lambda^k\big)\big\| = +\infty.$$

Since $\{\lambda^k\} \subset \Lambda$ and Λ is a compact set, we get that

$$\lim_{k\to\infty} \big\|x^k\big\| = +\infty.$$

For each k, we know that $H_\beta(x^k,\lambda^k,\xi) \in K$ by (7). Taking into account that the set K is compact, the sequence $\{H_\beta(x^k,\lambda^k,\xi)\}$ is also bounded. Therefore, we obtain that $\|x^k - H_\beta(x^k,\lambda^k,\xi)\| \to \infty$, as k tends to ∞.

It follows from the definition of $\theta(x,\lambda)$ and (15) in Lemma 6 that

$$
\begin{aligned}
\theta\big(x^k,\lambda^k\big) &= \mathbb{E}\big[g_{\alpha\beta}\big(x^k,\lambda^k,\xi\big)\big] \\
&\geq \mathbb{E}\left[\frac{\beta-\alpha}{2}\big\|x^k - H_\beta\big(x^k,\lambda^k,\xi\big)\big\|^2\right] \\
&= \frac{\beta-\alpha}{2}\mathbb{E}\big[\big\|x^k - H_\beta\big(x^k,\lambda^k,\xi\big)\big\|^2\big].
\end{aligned}
\tag{16}
$$

Since $\|x^k - H_\beta(x^k,\lambda^k,\xi)\| \to \infty$, as k tends to ∞, we get that $\mathbb{E}[\|x^k - H_\beta(x^k,\lambda^k,\xi)\|^2] \to \infty$, and so

$$\lim_{k\to\infty}\left(\frac{\beta-\alpha}{2}\mathbb{E}\big[\big\|x^k - H_\beta\big(x^k,\lambda^k,\xi\big)\big\|^2\big]\right) \to \infty,$$

whenever $0 < \alpha < \beta$. This, together with (16), implies that $\theta(x^k,\lambda^k) \to \infty$ as k tends to ∞. This contradicts the fact that $(x^k,\lambda^k) \in L_\theta(\bar{c})$, completing the proof. □

4 Convergence analysis

In this section, we utilize the sample average approximation (SAA) approach (for more about the SAA approach, please refer to, for example, [31]) when dealing with the expected value of the objective function.

Lemma 7 ([29]) *Let $\Phi : \Xi \to \mathbb{R}$ be an integrable function. Then, one has*

$$\mathbb{E}\big[\Phi(\xi)\big] = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N}\Phi(\xi_i), \quad \text{w.p.1}, \tag{17}$$

where $\{\xi_1, \xi_2, \dots, \xi_N\}$ is the independent and identical distributed (iid) sample of ξ and "w.p.1" denotes that this procedure converges with probability one.

Consequently, from Lemma 7, the UERM problem (8) is further converted into the following SAA problem:

$$\min_{x\in\mathbb{R}^n,\lambda\in\Lambda} \theta_N(x,\lambda) := \frac{1}{N}\sum_{i=1}^{N} g_{\alpha\beta}(x,\lambda,\xi_i). \tag{18}$$

4.1 Convergence of global optimal solutions

In this subsection, we will investigate the limiting behavior of global optimal solutions. We denote by S^* and S_N^* the sets of optimal solutions for problems (8) and (18), respectively.

Lemma 8 *For any $(x,\lambda) \in \mathbb{R}^n \times \Lambda$, one has*

$$\theta(x,\lambda) = \lim_{N\to\infty} \theta_N(x,\lambda), \quad \text{w.p.1}. \tag{19}$$

Proof Since the D-gap function $g_{\alpha\beta}(x,\lambda,\cdot)$ is integrable on Ξ, by Lemma 7, we obtain that

$$\lim_{N\to\infty} \theta_N(x,\lambda) = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} g_{\alpha\beta}(x,\lambda,\xi_i) = \mathbb{E}\big[g_{\alpha\beta}(x,\lambda,\xi)\big] = \theta(x,\lambda).$$

This completes the proof. □

Theorem 3 *Suppose that the assumptions* (a) *and* (b) *hold. Let $(x^N,\lambda^N) \in S_N^*$ for each N and let (x^*,λ^*) be an accumulation point of the sequence $\{(x^N,\lambda^N)\}$. Then, we have $(x^*,\lambda^*) \in S^*$.*

Proof Let (x^*,λ^*) be an accumulation point of the sequence $\{(x^N,\lambda^N)\}$. Without loss of generality, we assume that the sequence $\{(x^N,\lambda^N)\}$ converges to a point (x^*,λ^*). It is obvious that $(x^*,\lambda^*) \in \mathbb{R}^n \times \Lambda$. We divide the proof into two parts.

Part 1. We claim that

$$\lim_{N\to\infty} \big(\theta_N\big(x^N,\lambda^N\big) - \theta_N\big(x^*,\lambda^*\big)\big) = 0. \tag{20}$$

In fact, from Theorem 1, for any $(x,\lambda) \in \mathbb{R}^n \times \Lambda$ and $\xi \in \Xi$, we have

$$\big\|\nabla_x g_{\alpha\beta}(x,\lambda,\xi)\big\| \le \sum_{j=1}^{m}\big\|F_j(x,\xi)\big\|\left(4 + \big(\alpha^{-1} - \beta^{-1}\big)\sum_{j=1}^{m}\big\|\nabla_x F_j(x,\xi)\big\|\right) \tag{21}$$

and

$$\left\|\nabla_\lambda g_{\alpha\beta}(x,\lambda,\xi)\right\| \le \left(\alpha^{-1}+\beta^{-1}\right)\left(\sum_{j=1}^{m}\left\|F_j(x,\xi)\right\|\right)^2. \tag{22}$$

It follows from the mean-value theorem that for any each (x^N,λ^N) and each ξ_i,

$$\begin{aligned}
&\left|g_{\alpha\beta}\left(x^N,\lambda^N,\xi_i\right)-g_{\alpha\beta}\left(x^*,\lambda^*,\xi_i\right)\right| \\
&\quad = \left|\nabla_x g_{\alpha\beta}\left(y^{Ni},\lambda^{Ni},\xi_i\right)^\top\left(x^N-x^*\right)+\nabla_\lambda g_{\alpha\beta}\left(y^{Ni},\lambda^{Ni},\xi_i\right)^\top\left(\lambda^N-\lambda^*\right)\right|,
\end{aligned} \tag{23}$$

where $(y^{Ni},\lambda^{Ni}) \in \mathbb{R}^n \times \Lambda$ and $y^{Ni} = a_{Ni}x^N + (1-a_{Ni})x^*$, $\lambda^{Ni} = a_{Ni}\lambda^N + (1-a_{Ni})\lambda^*$ with $a_{Ni} \in [0,1]$. Then we get that

$$\begin{aligned}
&\left|\theta_N\left(x^N,\lambda^N\right)-\theta_N\left(x^*,\lambda^*\right)\right| \\
&\quad \le \frac{1}{N}\sum_{i=1}^{N}\left|g_{\alpha\beta}\left(x^N,\lambda^N,\xi_i\right)-g_{\alpha\beta}\left(x^*,\lambda^*,\xi_i\right)\right| \\
&\quad \le \frac{1}{N}\sum_{i=1}^{N}\left(\left\|\nabla_x g_{\alpha\beta}\left(y^{Ni},\lambda^{Ni},\xi_i\right)\right\|\left\|x^N-x^*\right\|+\left\|\nabla_\lambda g_{\alpha\beta}\left(y^{Ni},\lambda^{Ni},\xi_i\right)\right\|\left\|\lambda^N-\lambda^*\right\|\right) \\
&\quad \le \frac{1}{N}\sum_{i=1}^{N}\left(\sum_{j=1}^{m}\left\|F_j\left(y^{Ni},\xi_i\right)\right\|\right)\left(4+\left(\alpha^{-1}-\beta^{-1}\right)\sum_{j=1}^{m}\left\|\nabla_x F_j\left(y^{Ni},\xi_i\right)\right\|\right)\left\|x^N-x^*\right\| \\
&\qquad + \frac{1}{N}\sum_{i=1}^{N}\left(\alpha^{-1}+\beta^{-1}\right)\left(\sum_{j=1}^{m}\left\|F_j\left(y^{Ni},\xi_i\right)\right\|\right)^2\left\|\lambda^N-\lambda^*\right\| \\
&\quad \le \frac{1}{N}\sum_{i=1}^{N}\left(\sum_{j=1}^{m}\left\|F_j\left(y^{Ni},\xi_i\right)\right\|\right)\left(4+\left(\alpha^{-1}-\beta^{-1}\right)\sum_{j=1}^{m}\left\|\nabla_x F_j\left(y^{Ni},\xi_i\right)\right\|_{\mathcal{F}}\right)\left\|x^N-x^*\right\| \\
&\qquad + \frac{1}{N}\sum_{i=1}^{N}\left(\alpha^{-1}+\beta^{-1}\right)\left(\sum_{j=1}^{m}\left\|F_j\left(y^{Ni},\xi_i\right)\right\|\right)^2\left\|\lambda^N-\lambda^*\right\| \\
&\quad \le \frac{1}{N}\sum_{i=1}^{N}\kappa(\xi_i)\left(4+\left(\alpha^{-1}-\beta^{-1}\right)\kappa(\xi_i)\right)\left\|x^N-x^*\right\| \\
&\qquad + \left(\alpha^{-1}+\beta^{-1}\right)\frac{1}{N}\sum_{i=1}^{N}\kappa^2(\xi_i)\left\|\lambda^N-\lambda^*\right\| \\
&\quad \xrightarrow{N\to\infty} 0,
\end{aligned}$$

where the second inequality follows from (23) and Cauchy–Schwarz inequality, the third inequality follows from (21) and (22), the fourth inequality follows from the definition of the Frobenius matrix norm, and the last inequality follows from assumption (b). Since the sequence $\{(x^N,\lambda^N)\}$ converges to a point (x^*,λ^*), using assumption (b), we conclude that (20) is true.

Part 2. We next show that $(x^*,\lambda^*) \in S^*$. Since

$$\left|\theta_N\left(x^N,\lambda^N\right)-\theta\left(x^*,\lambda^*\right)\right| \le \left|\theta_N\left(x^N,\lambda^N\right)-\theta_N\left(x^*,\lambda^*\right)\right| + \left|\theta_N\left(x^*,\lambda^*\right)-\theta\left(x^*,\lambda^*\right)\right|,$$

it follows from Lemma 8 and (20) that

$$\lim_{N\to\infty} \theta_N\big(x^N,\lambda^N\big) = \theta\big(x^*,\lambda^*\big) \quad \text{w.p.1.}$$

Notice that $(x^N,\lambda^N) \in S_N^*$ for each N, which means that

$$\theta_N\big(x^N,\lambda^N\big) \le \theta_N(x,\lambda), \quad \forall (x,\lambda) \in \mathbb{R}^n \times \Lambda.$$

Taking the limit in the above inequality as $N \to \infty$, we obtain that

$$\theta\big(x^*,\lambda^*\big) \le \theta(x,\lambda), \quad \forall (x,\lambda) \in \mathbb{R}^n \times \Lambda \text{w.p.1,}$$

which implies that $(x^*,\lambda^*) \in S^*$ with probability one. □

4.2 Convergence of stationary points

A point (x^*,λ^*) is said to be a stationary point for problem (8) if it satisfies

$$\nabla\theta\big(x^*,\lambda^*\big) = 0. \tag{24}$$

For each N, a point (x^N,λ^N) is said to be a stationary point for problem (18) if it satisfies

$$\nabla\theta_N\big(x^N,\lambda^N\big) = 0. \tag{25}$$

Theorem 4 *Suppose that assumptions* (a), (b), *and* (c) *hold. Let* (x^N,λ^N) *be a stationary point of problem* (18) *for each* N *and* (x^*,λ^*) *be an accumulation point of the sequence* $\{(x^N,\lambda^N)\}$. *Then,* (x^*,λ^*) *is a stationary point of problem* (8) *with probability one.*

Proof In view of the definitions of $\theta(\cdot,\cdot)$ and $\theta_N(\cdot,\cdot)$, we have

$$\nabla\theta(x,\lambda) = \begin{pmatrix} \mathbb{E}[\sum_{j=1}^m \lambda_j \nabla_x F_j(x,\xi)(H_\beta(x,\lambda,\xi) - H_\alpha(x,\lambda,\xi))] + \alpha\mathbb{E}[(H_\alpha(x,\lambda,\xi) - x)] - \beta\mathbb{E}[(H_\beta(x,\lambda,\xi) - x)] \\ \mathbb{E}[(H_\beta(x,\lambda,\xi) - H_\alpha(x,\lambda,\xi))^\top F_1(x,\xi)] \\ \vdots \\ \mathbb{E}[(H_\beta(x,\lambda,\xi) - H_\alpha(x,\lambda,\xi))^\top F_m(x,\xi)] \end{pmatrix}$$

and

$$\nabla\theta_N(x,\lambda) = \begin{pmatrix} \frac{1}{N}\sum_{i=1}^N (\sum_{j=1}^m \lambda_j \nabla_x F_j(x,\xi_i)(H_\beta(x,\lambda,\xi_i) - H_\alpha(x,\lambda,\xi_i))) + \frac{\alpha}{N}\sum_{i=1}^N (H_\alpha(x,\lambda,\xi_i) - x) - \frac{\beta}{N}\sum_{i=1}^N (H_\beta(x,\lambda,\xi_i) - x) \\ \frac{1}{N}\sum_{i=1}^N ((H_\beta(x,\lambda,\xi_i) - H_\alpha(x,\lambda,\xi_i))^\top F_1(x,\xi_i)) \\ \vdots \\ \frac{1}{N}\sum_{i=1}^N ((H_\beta(x,\lambda,\xi_i) - H_\alpha(x,\lambda,\xi_i))^\top F_m(x,\xi_i)) \end{pmatrix}.$$

Let $D \subset \mathbb{R}^n \times \Lambda$ be a compact set. We have from assumptions (a), (b), and (c) that

$$\lim_{N\to\infty} \sup_{(x,\lambda)\in D} \big\|\nabla\theta_N(x,\lambda) - \nabla\theta(x,\lambda)\big\| = 0 \quad \text{w.p. 1.} \tag{26}$$

From assumptions (a) and (b), it is easy to see that $\nabla\theta(\cdot,\cdot)$ is a continuous function. Let (x^*,λ^*) be an accumulation point of the sequence $\{(x^N,\lambda^N)\}$. Without loss of generality, we assume that $\lim_{N\to\infty}(x^N,\lambda^N)=(x^*,\lambda^*)$. The sequence $\{(x^N,\lambda^N)\}$ is contained in a closed neighborhood $B\subset\mathbb{R}^n\times\Lambda$ of (x^*,λ^*) for sufficiently large N. Thus we conclude that for any given $\varepsilon>0$,

$$\left\|\nabla\theta\left(x^N,\lambda^N\right)-\nabla\theta\left(x^*,\lambda^*\right)\right\|\leq\frac{\varepsilon}{2}. \tag{27}$$

From (26), there exists $N_0>0$ such that $(x^N,\lambda^N)\in B$ for all $N\geq N_0$ and

$$\left\|\nabla\theta_N\left(x^N,\lambda^N\right)-\nabla\theta\left(x^N,\lambda^N\right)\right\|\leq\frac{\varepsilon}{2}. \tag{28}$$

Since

$$\begin{aligned}&\left\|\nabla\theta_N\left(x^N,\lambda^N\right)-\nabla\theta\left(x^*,\lambda^*\right)\right\|\\&\quad\leq\left\|\nabla\theta_N\left(x^N,\lambda^N\right)-\nabla\theta\left(x^N,\lambda^N\right)\right\|+\left\|\nabla\theta\left(x^N,\lambda^N\right)-\nabla\theta\left(x^*,\lambda^*\right)\right\|,\end{aligned}$$

we have from (27) and (28) that

$$\left\|\nabla\theta_N\left(x^N,\lambda^N\right)-\nabla\theta\left(x^*,\lambda^*\right)\right\|\leq\varepsilon.$$

Thus, we obtain that

$$\lim_{N\to\infty}\nabla\theta_N\left(x^N,\lambda^N\right)=\nabla\theta\left(x^*,\lambda^*\right)\quad\text{w.p. 1.}$$

By taking the limit as N tends to ∞ in (25), we obtain (24). That is, (x^*,λ^*) is stationary point of problem (8) with probability one. This completes the proof. □

5 EV formulation and its global error bound

In this section, let us consider another deterministic formulation for the SVVI problem, i.e., the EV formulation (for more details about the EV formulation, please see, for example, [11]), that is, to find $x^*\in K$ such that

$$\begin{aligned}&\left(\left(y-x^*\right)^\top\mathbb{E}\left[F_1\left(x^*,\xi\right)\right],\dots,\left(y-x^*\right)^\top\mathbb{E}\left[F_m\left(x^*,\xi\right)\right]\right)\\&\quad\notin\text{-int}\,\mathbb{R}^m_+,\quad\forall y\in K,\text{a.s. }\xi\in\Xi.\end{aligned} \tag{29}$$

The notations $\text{sol}(\mathbb{E}[F(x,\xi)],K)$ denotes the solution set of problem (29) and $\text{dist}(x,X)$ denotes $\min_{y\in X}\|x-y\|$. Lee et al. [16] provided the following properties for the deterministic VVI.

Lemma 9 ([16, Theorem 4.2]) *Suppose that F_j are strongly monotone on K with modulus $\beta>0$ and Lipschitz continuous on K with modulus $l>0$ for all j $(j=1,\dots,m)$. Then the solution set is compact.*

Based on the EV formulation, we give some conditions in this section. Then, the D-gap function provides a global error bound for the SVVI problem (29). Assume that the expected value of the function $F_j(x,\cdot)$ is well defined. Similar to the work of Zhao et al. [33, Eq. (6)], we give an equivalent scalar variational inequality for problem (29), i.e., find $(x^*,\lambda^*) \in K \times \Lambda$ such that

$$\left(y - x^*\right)^\top \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\left[F_j\left(x^*, \xi\right)\right] \geq 0, \quad \forall y \in K. \tag{30}$$

From the ideas of Yamashita–Taji–Fukushima [32, Eq. (6)] and Zhao et al. [33, Eq. (7)], the D-gap function $g_{\alpha\beta}(\cdot,\cdot) : \mathbb{R}^n \times \Lambda$ of problem (30) is defined as follows:

$$g_{\alpha\beta}(x,\lambda) := g_\alpha(x,\lambda) - g_\beta(x,\lambda),$$

where $0 < \alpha < \beta$ and $g_\gamma(\cdot,\cdot)$ ($\gamma = \alpha, \beta$) is the regularized gap function, which is defined by

$$g_\gamma(x,\lambda) := \max_{y\in K} \left\{ (x-y)^\top \sum_{j=1}^{m} \lambda_j \mathbb{E}\left[F_j(x,\xi)\right] - \frac{\gamma}{2}\|x-y\|^2 \right\}.$$

It is easy to see that

$$g_\gamma(x,\lambda) := \left(x - H_\gamma(x,\lambda)\right)^\top \sum_{j=1}^{m} \lambda_j \mathbb{E}\left[F_j(x,\xi)\right] - \frac{\gamma}{2}\left\|x - H_\gamma(x,\lambda)\right\|^2, \tag{31}$$

where

$$H_\gamma(x,\lambda) := \mathrm{Proj}_K\left(x - \gamma^{-1}\sum_{j=1}^{m} \lambda_j \mathbb{E}\left[F_j(x,\xi)\right]\right). \tag{32}$$

The following main properties of the D-gap function $g_{\alpha\beta}(\cdot,\cdot)$ hold:

- $g_{\alpha\beta}(x,\lambda) \geq 0$ for any $(x,\lambda) \in \mathbb{R}^n \times \Lambda$;
- $g_{\alpha\beta}(x^*,\lambda^*) = 0$ if and only if (x^*,λ^*) solves problem (30).

Remark 3 If $m = 1$, then the D-gap function $g_{\alpha\beta}(\cdot,\cdot)$ considered in the present paper reduces to the D-gap function [12, Eq. (5)] discussed by He et al.

As usual, $\mathbb{P}(V)$ denotes the probability of an event V.

Lemma 10 *Let* $0 < \alpha < \beta$. *For any* $(x,\lambda) \in \mathbb{R}^n \times \Lambda$, *one has*

$$\frac{\beta-\alpha}{2}\left\|x - H_\beta(x,\lambda)\right\|^2 \leq g_{\alpha\beta}(x,\lambda) \leq \frac{\beta-\alpha}{2}\left\|x - H_\alpha(x,\lambda)\right\|^2. \tag{33}$$

Proof This proof is similar to that of Lemma 6, so we omit it here. □

Lemma 11 *Suppose that assumptions* (a) *and* (c) *hold, and each function* $F_j(\cdot,\xi)$ ($j = 1,\dots,m$) *is monotone on* K *for almost every* $\xi \in \Xi$. *Let* F_j ($j = 1,\dots,m$) *be uniformly strongly monotone on* K *with modulus* $\mu_j > 0$ *over* $V_j \subset \Xi \subset \mathbb{R}^n$ *with* $\mathbb{P}(V_j) > 0$, $\mu := \min_{1\leq j\leq m}\mu_j$ *and* $\nu := \min_{1\leq j\leq m}\mathbb{P}(V_j)$. *Then the solution set* $\mathrm{sol}(\mathbb{E}[F(x,\xi)],K)$ *is compact.*

Proof For each $j = 1, \dots, m$, since F_j is uniformly strongly monotone on K with modulus $\mu_j > 0$ over V_j, we have

$$\begin{aligned} \int_{V_j} \bigl(F_j(y,\xi) - F_j(x,\xi)\bigr)^\top (y-x)\mathbb{P}(d\xi) &\geq \int_{V_j} \mu_j \|y-x\|^2 \mathbb{P}(d\xi) \\ &= \mu_j \|y-x\|^2 \int_{V_j} \mathbb{P}(d\xi) \\ &= \mu_j \|y-x\|^2 \mathbb{P}(V_j) \\ &\geq \mu\nu \|y-x\|^2. \end{aligned} \tag{34}$$

In addition, since $F_j(\cdot,\xi)$ is monotone on $\mathbb{R}^n$ for almost every $\xi \in \Xi$, we have

$$\int_{\Xi\setminus V_j} \bigl(F_j(y,\xi) - F_j(x,\xi)\bigr)^\top (y-x)\mathbb{P}(d\xi) \geq \int_{\Xi\setminus V_j} 0\mathbb{P}(d\xi) = 0. \tag{35}$$

Combining (34) and (35), for any $x, y \in \mathbb{R}^n$, we get

$$\begin{aligned} \bigl(\mathbb{E}\bigl[F_j(y,\xi)\bigr] - \mathbb{E}\bigl[F_j(x,\xi)\bigr]\bigr)^\top (y-x) &= \int_{V_j} \bigl(F_j(y,\xi) - F_j(x,\xi)\bigr)^\top (y-x)\mathbb{P}(d\xi) \\ &\quad + \int_{\Xi\setminus V_j} \bigl(F_j(y,\xi) - F_j(x,\xi)\bigr)^\top (y-x)\mathbb{P}(d\xi) \\ &\geq \nu\mu \|y-x\|^2. \end{aligned}$$

Hence, for each j ($j = 1, \dots, m$), $\mathbb{E}[F_j]$ is strongly monotone with modulus $\nu\mu > 0$. On the other hand, by assumption (c), for any $x, y \in \mathbb{R}^n$, it holds that

$$\begin{aligned} \bigl\|\mathbb{E}\bigl[F_j(y,\xi)\bigr] - \mathbb{E}\bigl[F_j(x,\xi)\bigr]\bigr\| &= \int_{\Xi} \bigl\|F_j(y,\xi) - F_j(x,\xi)\bigr\| \mathbb{P}(d\xi) \\ &\leq \int_{\Xi} L_j(\xi)\|y-x\|\mathbb{P}(d\xi) \\ &\leq L\|y-x\|. \end{aligned}$$

Therefore, for each j ($j = 1, \dots, m$), $\mathbb{E}[F_j]$ is Lipschitz continuous with modulus $L > 0$ and strongly monotone with modulus $\nu\mu > 0$. Then, from Lemma 9, we obtain that $\mathrm{sol}(\mathbb{E}[F(x,\xi)], K)$ is compact. □

Theorem 5 *Suppose that assumptions* (a), (b), *and* (c) *hold, and each function* $F_j(\cdot,\xi)$ ($j = 1, \dots, m$) *is monotone on* K *for almost every* $\xi \in \Xi$. *Let* F_j ($j = 1, \dots, m$) *be uniformly strongly monotone on* K *with modulus* $\mu_j > 0$ *over* $V_j \subset \Xi \subset \mathbb{R}^n$ *with* $\mathbb{P}(V_j) > 0$, $\mu := \min_{1\leq j\leq m} \mu_j$ *and* $\nu := \min_{1\leq j\leq m} \mathbb{P}(V_j)$. *Thus, for any* $\lambda \in \Lambda$, *one has*

$$\mathrm{dist}\bigl(x, \mathrm{sol}\bigl(\mathbb{E}\bigl[F(x,\xi)\bigr], K\bigr)\bigr) \leq \frac{L+\beta}{\nu\mu}\sqrt{\frac{2}{\beta-\alpha}}\sqrt{g_{\alpha\beta}(x,\lambda)}.$$

Proof From (30), it is natural that as we change $\lambda \in \Lambda$, x will also change. Furthermore, once we have fixed $\lambda^* \in \Lambda$, the corresponding x is also fixed. Since each function F_j ($j =$

$1, \dots, m$) is uniformly strongly monotone on $\mathbb{R}^n$, using the concepts of monotonicity and strong monotonicity, for any $x, y \in \mathbb{R}^n$, one has

$$\begin{aligned}
&\left(\sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j(y,\xi)\big] - \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j(x,\xi)\big]\right)^\top (y-x) \\
&\quad = \sum_{j=1}^{m} \int_{V_j} \lambda_j^* \big(F_j(y,\xi) - F_j(x,\xi)\big)^\top (y-x)\mathbb{P}(d\xi) \\
&\qquad + \sum_{j=1}^{m} \int_{\Xi\setminus V_j} \lambda_j^* \big(F_j(y,\xi) - F_j(x,\xi)\big)^\top (y-x)\mathbb{P}(d\xi) \\
&\quad \ge \nu\mu \|y-x\|^2,
\end{aligned} \tag{36}$$

where $\lambda_j^* \in [0,1]$. Hence, $\sum_{j=1}^m \lambda_j^* \mathbb{E}[F_j]$ is strongly monotone with modulus $\nu\mu > 0$ for fixed $\lambda^* \in \Lambda$. Therefore, from Lemma 1, there is a unique solution x^* such that

$$\big(y - x^*\big)^\top \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j\big(x^*,\xi\big)\big] \ge 0, \quad \forall y \in K. \tag{37}$$

For any $x, y \in \mathbb{R}^n$, using the concepts of monotonicity and Lipschitz continuity, we obtain

$$\begin{aligned}
&\left(\sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j(y,\xi)\big] - \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j(x,\xi)\big]\right)^\top (y-x) \\
&\quad = \sum_{j=1}^{m} \int_{\Xi} \lambda_j^* \big(F_j(y,\xi) - F_j(x,\xi)\big)^\top (y-x)\mathbb{P}(d\xi) \\
&\quad \le L\|y-x\|^2,
\end{aligned} \tag{38}$$

where $\lambda_j^* \in [0,1]$. Hence, $\sum_{j=1}^m \lambda_j^* \mathbb{E}[F_j]$ is Lipschitz continuous with modulus $L > 0$. Then, we claim that there exists a constant $r > 0$ such that

$$\operatorname{dist}\big(x, \operatorname{sol}\big(\mathbb{E}\big[F(x,\xi)\big], K\big)\big) \le \big\|x - x^*\big\| \le r\big\|H_\beta\big(x,\lambda^*\big) - x\big\|. \tag{39}$$

In fact, given $x \in K$, from item (a) of Lemma 2, we know that $H_\beta(x,\lambda^*)$, defined by (32), is the unique solution of the following strongly convex minimization problem:

$$\min_{y\in K} \left\{ \left\langle \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j(x,\xi)\big], y - x \right\rangle + \frac{\beta}{2}\|y-x\|^2 \right\}.$$

Hence, $H_\beta(x,\lambda^*)$ fulfills, for all $y \in K$, the following optimality condition:

$$\left\langle \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j(x,\xi)\big] - \beta\big(x - H_\beta\big(x,\lambda^*\big)\big), y - H_\beta\big(x,\lambda^*\big) \right\rangle \ge 0.$$

Taking $y = x^*$ in the above inequality, we obtain

$$\left\langle \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j(x,\xi)\big] - \beta\big(x - H_\beta\big(x,\lambda^*\big)\big), H_\beta\big(x,\lambda^*\big) - x^* \right\rangle \le 0. \tag{40}$$

From (36), the function $\sum_{j=1}^{m} \lambda_j^* \mathbb{E}[F_j(\cdot,\xi)]$ is strongly monotone with modulus $\mu\nu$. Hence, there exists a unique solution x^* such that

$$\left\langle \sum_{j=1}^{m} \lambda_j \mathbb{E}\big[F_j\big(x^*,\xi\big)\big], x^* - H_\beta\big(x,\lambda^*\big) \right\rangle \le 0. \tag{41}$$

Adding (40) with (41), we have

$$\left\langle \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j(x,\xi)\big] - \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j\big(x^*,\xi\big)\big] - \beta\big(x - H_\beta\big(x,\lambda^*\big)\big), H_\beta\big(x,\lambda^*\big) - x^* \right\rangle \le 0.$$

This can be rewritten as follows:

$$\begin{aligned}
&\left\langle \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j(x,\xi)\big] - \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j\big(x^*,\xi\big)\big], x - x^* \right\rangle \\
&\quad \le -\beta \big\| H_\beta\big(x,\lambda^*\big) - x \big\|^2 \\
&\qquad - \left\langle \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j(x,\xi)\big] - \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j\big(x^*,\xi\big)\big], H_\beta\big(x,\lambda^*\big) - x \right\rangle \\
&\qquad + \beta \big\langle H_\beta\big(x,\lambda^*\big) - x, x^* - x \big\rangle.
\end{aligned} \tag{42}$$

Noticing that $\beta > 0$ and simultaneously utilizing Cauchy–Schwarz inequality, we get that

$$\begin{aligned}
&\left\langle \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j(x,\xi)\big] - \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j\big(x^*,\xi\big)\big], x - x^* \right\rangle \\
&\quad \le \left\| \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j(x,\xi)\big] - \sum_{j=1}^{m} \lambda_j^* \mathbb{E}\big[F_j\big(x^*,\xi\big)\big] \right\| \big\| H_\beta\big(x,\lambda^*\big) - x \big\| \\
&\qquad + \beta \big\| H_\beta\big(x,\lambda^*\big) - x \big\| \big\| x - x^* \big\|.
\end{aligned}$$

This, together with (36) and (38), implies that

$$\nu\mu \big\| x - x^* \big\|^2 \le L \big\| x - x^* \big\| \big\| H_\beta\big(x,\lambda^*\big) - x \big\| + \beta \big\| H_\beta\big(x,\lambda^*\big) - x \big\| \big\| x - x^* \big\|,$$

which means that

$$\operatorname{dist}\big(x, \operatorname{sol}\big(\mathbb{E}\big[F(x,\xi)\big], K\big)\big) \le \big\| x - x^* \big\| \le \frac{L + \beta}{\nu\mu} \big\| H_\beta\big(x,\lambda^*\big) - x \big\|. \tag{43}$$

As a result, the needed constant $r > 0$ is defined by $r = \frac{L+\beta}{\nu\mu}$. Thus, (39) is true. From the first inequality of (33) in Lemma 10, we have

$$\left\| x - H_\beta\left(x, \lambda^*\right) \right\| \le \sqrt{\frac{2}{\beta - \alpha}} \sqrt{g_{\alpha\beta}\left(x, \lambda^*\right)}. \tag{44}$$

Combining (43) with (44) and using the arbitrariness of $\lambda^* \in \Lambda$, we get the desired conclusion. This completes the proof. □

In the rest of this section, we investigate the boundedness of the level set of the D-gap function $g_{\alpha\beta}(\cdot,\cdot)$. The level set of the D-gap function $g_{\alpha\beta}(\cdot,\cdot)$ is defined by

$$L_{g_{\alpha\beta}}(\eta) := \left\{(x, \lambda) \in \mathbb{R}^n \times \Lambda : g_{\alpha\beta}(x, \lambda) \le \eta\right\}.$$

Corollary 1 *Suppose that assumptions* (a), (b), *and* (c) *hold, and each function* $F_j(\cdot, \xi)$ $(j = 1, \dots, m)$ *is monotone on* K *for almost every* $\xi \in \Xi$. *Let* F_j $(j = 1, \dots, m)$ *be uniformly strongly monotone on* K *with modulus* $\mu_j > 0$ *over* $V_j \subset \Xi \subset \mathbb{R}^n$ *with* $\mathbb{P}(V_j) > 0$, $\mu := \min_{1\le j\le m} \mu_j$ *and* $\nu := \min_{1\le j\le m} \mathbb{P}(V_j)$. *Then, for any* $\eta \ge 0$, *the level set* $L_{g_{\alpha\beta}}(\eta)$ *is bounded.*

Proof Suppose on the contrary that there exists a $\bar{\eta} \ge 0$ such that $L_{g_{\alpha\beta}}(\bar{\eta})$ is unbounded. This implies that there exists a sequence $\{(x^k, \lambda^k)\} \subset L_{g_{\alpha\beta}}(\bar{\eta})$ such that

$$\lim_{k\to\infty} \left\| \left(x^k, \lambda^k\right) \right\| = +\infty.$$

Since $\{\lambda^k\} \subset \Lambda$ and Λ is a compact set, we get that

$$\lim_{k\to\infty} \left\| x^k \right\| = +\infty.$$

By the proof of Theorem 5, one has

$$\operatorname{dist}\left(x^k, \operatorname{sol}\left(\mathbb{E}\left[F\left(x^k, \xi\right)\right], K\right)\right) \le \frac{L+\beta}{\nu\mu} \sqrt{\frac{2}{\beta - \alpha}} \sqrt{g_{\alpha\beta}\left(x^k, \lambda^k\right)}.$$

This means that

$$g_{\alpha\beta}\left(x^k, \lambda^k\right) \ge \frac{\nu^2\mu^2(\beta - \alpha)}{2(L+\beta)^2} \operatorname{dist}^2\left(x^k, \operatorname{sol}\left(\mathbb{E}\left[F\left(x^k, \xi\right)\right], K\right)\right).$$

Then, taking the limit of the right-hand side of the above inequality, we have

$$\lim_{k\to\infty} \left(\frac{\nu^2\mu^2(\beta - \alpha)}{2(L+\beta)^2} \operatorname{dist}^2\left(x^k, \operatorname{sol}\left(\mathbb{E}\left[F\left(x^k, \xi\right)\right], K\right)\right) \right) = \infty,$$

which means that $g_{\alpha\beta}(x^k, \lambda^k) \to \infty$ as k tends to ∞. This contradicts the fact that $(x^k, \lambda^k) \in L_{g_{\alpha\beta}}(\bar{\eta})$, completing the proof. □

Remark 4 We would like to point out that the condition that each function F_j $(j = 1, \dots, m)$ is Lipschitz continuous with Lipschitz constant $L_j(\xi)$ used in Theorem 5 and Corollary 1 can be replaced by a requirement that K is a compact set.

We only prove the corresponding result of Theorem 5. In fact, since K is a compact set, there exists a positive number $b > 0$ such that $K \subset \{x \mid \|x\| \leq b\}$. Let $\mathfrak{B}$ be a closed ball with radius $3b$, that is, $\mathfrak{B} = \{x \mid \|x\| \leq 3b\}$. For any $x \in \mathbb{R}^n$, let us consider two possible cases:

(Case (i): $x \in \mathfrak{B}$) Since each function $F_j(\cdot, \xi)$ $(j = 1, \dots, m)$ is continuously differentiable, each function $F_j(\cdot, \xi)$ $(j = 1, \dots, m)$ is Lipschitz continuous on $\mathfrak{B}$. From the proof of Theorem 5, for all $x \in \mathfrak{B}$, we have $\|x - x^*\| \leq \frac{L+\beta}{\nu\mu}\|H_\beta(x, \lambda^*) - x\|$.

(Case (ii): $x \notin \mathfrak{B}$) For an arbitrary $x \notin \mathfrak{B}$, we have $\|x\| \geq 3b$. Since $x^* \in K$, $H_\beta(x, \lambda^*) \in K$ and K is a compact set, we obtain that $\|x^*\| \leq b$ and $\|H_\beta(x, \lambda^*)\| \leq b$. Hence, we have

$$\|x - x^*\| \leq \|x - H_\beta(x, \lambda^*)\| + \|H_\beta(x, \lambda^*)\| + \|x^*\| \leq \|x - H_\beta(x, \lambda^*)\| + 2b. \tag{45}$$

It follows from the triangle inequality that

$$\|x - H_\beta(x, \lambda^*)\| \geq \|x\| - \|H_\beta(x, \lambda^*)\| \geq 3b - b = 2b.$$

This, together with (45), implies that $\|x - x^*\| \leq 2\|x - H_\beta(x, \lambda^*)\|$.

Setting $\bar{r} = \max\{\frac{L+\beta}{\nu\mu}, 2\}$, we have $\|x - x^*\| \leq \bar{r}\|x - H_\beta(x, \lambda^*)\|$, which plays the role of (39). Then by repeating the rest of the proof of Theorem 5, we get the desired conclusion.

Similarly, the analogue for Corollary 1 is true in the case where the set K is compact.

6 Concluding remarks

In the present paper, mainly motivated by the works [20, 32, 33], we have presented an UERM approach (i.e., problem (8)) for solving the SVVI problem (1). Several properties of the objective function θ were discussed, namely, the continuous differentibility and the boundedness of the level set. Furthermore, a well-known sample average approximation approach was presented for solving problem (8). The convergence of the proposed approach for global optimal solutions and stationary points was analyzed. Finally, we considered another deterministic formulation, i.e., the EV formulation for the SVVI problem. And the global error bound of the D-gap function $g_{\alpha\beta}(\cdot, \cdot)$ based on the EV formulation was given.

We would like to mention that the UERM approach presented in this paper is formalistic, since the projection operator on K is still needed in the calculation of $g_{\alpha\beta}$. Thus, in order to convert the formalistic approach into practical methods, it is interesting to investigate the following aspects:

(i) Based on the structure of the constraint set K, how to design effective algorithms;
(ii) The sample complexity of the SAA approach proposed in this paper;
(iii) The convergence rate of the SAA approach.

Acknowledgements
The authors are grateful to the editor and referee for their valuable comments and suggestions.

Funding
This work was supported by National Natural Science Foundation of China (No. 11961006), Guangxi Natural Science Foundation (2020GXNSFAA159100) and Innovation Project of Guangxi Graduate Education (gxun-chxs2021055).

Availability of data and materials
Not applicable.

Declarations

Competing interests
The authors declare no competing interests.

Author contributions
Dan-Dan Dong contributed to methodology and writing-original draft; Guo-ji Tang contributed to conceptualization, methodology and supervision; Hui-ming Qiu contributed to the revision of the manuscript.

Author details
[1]School of Mathematics and Physics, Guangxi Minzu University, Guangxi, 530006, P.R. China. [2]School of Mathematics and Physics, Guangxi Key Laboratory of Hybrid Computation and IC Design Analysis, Guangxi Minzu University, Guangxi, 530006, P.R. China.

Received: 1 April 2023 Accepted: 12 July 2023 Published online: 27 July 2023

References
1. Ansari, Q.H., Köbis, E., Yao, J.C.: Vector Variational Inequalities and Vector Optimization: Theory and Applications. Springer, Switzerland (2018)
2. Bianchi, M., Konnov, I.V., Pini, R.: Limit vector variational inequalities and market equilibrium problems. Optim. Lett. **15**, 817–832 (2021)
3. Charitha, C., Dutta, J.: Regularized gap functions and error bounds for vector variational inequalities. Pac. J. Optim. **6**, 497–510 (2010)
4. Charitha, C., Dutta, J., Lalitha, C.S.: Gap functions for vector variational inequalities. Optimization **64**, 1499–1520 (2015)
5. Chen, X.J., Fukushima, M.: Expected residual minimization method for stochastic linear complementarity problems. Math. Oper. Res. **30**, 1022–1038 (2005)
6. Daniele, P., Maugeri, A.: Vector variational inequalities and modelling of a continuum traffic equilibrium problem. In: Giannessi, F. (ed.) Vector Variational Inequalities and Vector Equilibria, Mathematical Theories, pp. 97–111. Kluwer Academic, Dordrecht (2000)
7. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume I. Springer, New York (2003)
8. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, vol. II. Springer, New York (2003)
9. Giannessi, F.: Theorems of alternative, quadratic programs and complementarity problems. In: Cottle, R.W., Giannessi, F., Lions, J.L. (eds.) Variational Inequalities and Complementarity Problems, pp. 151–186. Wiley, New York (1980)
10. Giannessi, F. (ed.): Vector Variational Inequalities and Vector Equilibria: Mathematical Theories. Nonconvex Optimization and Its Applications, vol. 38. Kluwer Academic, Dordrecht (2000)
11. Gürkan, G., Özge, A.Y., Robinson, S.M.: Sample-path solution of stochastic variational inequalities. Math. Program. **84**, 313–333 (1999)
12. He, S.X., Zhang, P., Hu, X., Hu, R.: A sample average approximation method based on a D-gap function for stochastic variational inequality problems. J. Ind. Manag. Optim. **10**, 977–987 (2014)
13. Huang, N.J., Li, J., Yang, X.Q.: Weak sharpness for gap functions in vector variational inequalities. J. Math. Anal. Appl. **394**, 449–457 (2012)
14. Jiang, H.Y., Xu, H.F.: Stochastic approximation approaches to the stochastic variational inequality problem. IEEE Trans. Autom. Control **53**, 1462–1475 (2008)
15. Kanzow, C., Fukushima, M.: Theoretical and numerical investigation of the D-gap function for box constrained variational inequalities. Math. Program. **83**, 55–87 (1998)
16. Lee, G.M., Kim, D.S., Lee, B.S., Yen, N.D.: Vector variational inequality as a tool for studying vector optimization problems. Nonlinear Anal. **34**, 745–765 (1998)
17. Lee, G.M., Yen, N.D.: A result on vector variational inequalities with polyhedral constraint sets. J. Optim. Theory Appl. **109**, 193–197 (2001)
18. Li, S.J., Chen, G.Y.: On relations between multiclass, multicriteria traffic network equilibrium models and vector variational inequalities. J. Syst. Sci. Syst. Eng. **15**, 284–297 (2006)
19. Lin, G.H., Fukushima, M.: Stochastic equilibrium problems and stochastic mathematical programs with equilibrium constraints: a survey. Pac. J. Optim. **6**, 455–482 (2010)
20. Liu, J.X., Li, S.J.: Unconstrained optimization reformulation for stochastic nonlinear complementarity problems. Appl. Anal. **100**, 1158–1179 (2021)
21. Lu, F., Li, S.J.: Method of weighted expected residual for solving stochastic variational inequality problems. Appl. Math. Comput. **269**, 651–663 (2015)
22. Lu, F., Li, S.J., Yang, J.: Convergence analysis of weighted expected residual method for nonlinear stochastic variational inequality problems. Math. Methods Oper. Res. **82**, 229–242 (2015)
23. Luo, M.J., Lin, G.H.: Expected residual minimization method for stochastic variational inequality problems. J. Optim. Theory Appl. **140**, 103–116 (2009)
24. Luo, M.J., Lin, G.H.: Convergence results of ERM method for nonlinear stochastic variational inequality problems. J. Optim. Theory Appl. **142**, 569–581 (2009)
25. Luo, M.J., Lin, G.H.: Stochastic variational inequality problems with additional constraints and their applications in supply chain network equilibria. Pac. J. Optim. **7**, 263–279 (2011)
26. Luo, M.J., Zhang, K.: Convergence analysis of the approximation problems for solving stochastic vector variational inequality problems. Complexity **2020**, Article ID 1203627 (2020)
27. Ma, H.Q., Huang, N.J., Wu, M., Regan, D.O.: A new gap function for vector variational inequalities with an application. J. Appl. Math. **2013**, Article ID 423040 (2013)
28. Ma, H.Q., Wu, M., Huang, N.J.: Expected residual minimization method for stochastic variational inequality problems with nonlinear perturbations. Appl. Math. Comput. **219**, 6256–6267 (2013)

29. Niederreiter, H.: Random Number Generation and Quasi-Monte Carlo Methods. SIAM, Philadelphia (1992)
30. Patrick, B.: Probability and Measure. Wiley, New York (1995)
31. Shapiro, A., Xu, H.F.: Stochastic mathematical programs with equilibrium constraints, modelling and sample average approximation. Optimization **57**, 395–418 (2008)
32. Yamashita, N., Taji, K., Fukushima, M.: Unconstrained optimization reformulations of variational inequality problems. J. Optim. Theory Appl. **92**, 439–456 (1997)
33. Zhao, Y., Zhang, J., Yang, X.M., Lin, G.H.: Expected residual minimization formulation for a class of stochastic vector variational inequalities. J. Optim. Theory Appl. **175**, 545–566 (2017)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.